Q-Learning for Robot Control

A thesis submitted for the degree of
Doctor of Philosophy
of The Australian National University.

Chris Gaskett
Bachelor of Computer Systems Engineering H1 (RMIT University)
Bachelor of Computer Science (RMIT University)

Supervisor:
Professor Alexander Zelinsky

2002

Statement of Originality

These doctoral studies were conducted under the supervision of Professor Alexander Zelinsky. The work submitted in this thesis is a result of original research carried out by myself, in collaboration with others, while enrolled as a PhD student in the Department of Systems Engineering at the Australian National University. It has not been submitted for any other degree or award.

Chris Gaskett

Abstract

Q-Learning is a method for solving reinforcement learning problems. Reinforcement learning problems require improvement of behaviour based on received rewards. Q-Learning has the potential to reduce robot programming effort and increase the range of robot abilities. However, most current Q-learning systems are not suitable for robotics problems: they treat continuous variables, for example speeds or positions, as discretised values. Discretisation does not allow smooth control and does not fully exploit sensed information. A practical algorithm must also cope with real-time constraints, sensing and actuation delays, and incorrect sensor data.

This research describes an algorithm that deals with continuous state and action variables without discretising. The algorithm is evaluated with vision-based mobile robot and active head gaze control tasks. As well as learning the basic control tasks, the algorithm learns to compensate for delays in sensing and actuation by predicting the behaviour of its environment. Although the learned dynamic model is implicit in the controller, it is possible to extract some aspects of the model. The extracted models are compared to theoretically derived models of environment behaviour.

The difficulty of working with robots motivates development of methods that reduce experimentation time. This research exploits Q-learning's ability to learn by passively observing the robot's actions—rather than necessarily controlling the robot. This is a valuable tool for shortening the duration of learning experiments.

Acknowledgements

I thank my supervisor, Professor Alexander Zelinsky, and my previous supervisor Dr David Wettergreen, for guiding me through this research. Thank you to all the staff and students of Systems Engineering, past and present, for their constant support and friendship. Thanks also to all of the visitors to the Robotic Systems Laboratory for their valuable advice.

The mobile robot software was built upon the work of Dr Gordon Cheng (1996), and used correlation software developed by Dr Jochen Heinzmann (1997). Luke Fletcher ported Dr Cheng's software to the Nomad, integrated the learning system into the software, participated in the experiments, helped with video editing, and co-authored two papers. The active head software and experiments were the work of myself and Peter Brown. Vision libraries were provided by Dr Sebastien Rougeaux. Orson Sutherland, Leanne Matuszyk, and Harley Truong provided active head system software. Dr Thomas Brinsmead and Peter Brown assisted with the theoretical pendulum model. Dr Simon Thompson assisted with video editing. Karen Montefiore, Gareth Loy, Luke Fletcher, and Dr Simon Thompson assisted with preparation of diagrams. Dr Peter Bartlett and Dr Jonathan Baxter gave valuable advice about learning systems theory. Dr Baxter also supplied some example feedforward neural network source code. Professor Brian Anderson, Professor John Moore, Professor Robert Bitmead, Dr Thomas Brinsmead, and Dr Thomas Moor answered my questions about control systems. The display software for the cart-pole task was based on Professor Charles Anderson's (1998) pole balancing computer game. Dr Gordon Cheng reviewed an earlier draft. James Ashton kept the computers' wheels turning. Jenni Watkins, Marita Rendina, Karen Montefiore, and Rosemary Shepherd arranged everything and provided moral support. Dr Jochen Heinzmann, Dr David Jung, Dr Gordon Cheng, Dr Rochelle O'Hagan, and Andrew Brooks made me feel welcome from the start. My friends from RMIT University taught me the value of teamwork done well.

My family gave me unflagging support throughout the preparation of this thesis. Finally, I thank my wife Rika, for everything.

Chris Gaskett

List of Publications

Gaskett, C., Brown, P., Cheng, G., and Zelinsky, A. (2003), Learning implicit models during target pursuit, *in Proceedings of the IEEE International Conference on Robotics and Automation (ICRA2003)*, Taiwan.

Gaskett, C., Fletcher, L., and Zelinsky, A. (2000a), Reinforcement learning for a vision based mobile robot, *in Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS2000)*, Takamatsu, Japan.

Gaskett, C., Fletcher, L., and Zelinsky, A. (2000b), Reinforcement learning for visual servoing of a mobile robot, *in Proceedings of the Australian Conference on Robotics and Automation (ACRA2000)*, Melbourne, Australia.

Gaskett, C., Wettergreen, D., and Zelinsky, A. (1999a), Q-learning in continuous state and action spaces, *in Proceedings of the 12th Australian Joint Conference on Artificial Intelligence*, Sydney, Australia.

Gaskett, C., Wettergreen, D., and Zelinsky, A. (1999b), Reinforcement learning applied to the control of an autonomous underwater vehicle, *in Proceedings of the Australian Conference on Robotics and Automation (AuCRA'99)*, Brisbane, Australia.

Wettergreen, D., Gaskett, C., and Zelinsky, A. (1998), Development of a visually-guided autonomous underwater vehicle, *in Proceedings of OCEANS98*, IEEE, Nice.

Wettergreen, D., Gaskett, C., and Zelinsky, A. (1999a), Autonomous control and guidance for an underwater robotic vehicle, *in Proceedings of the International Conference on Field and Service Robotics (FSR'99)*, Pittsburgh, USA.

Wettergreen, D., Gaskett, C., and Zelinsky, A. (1999b), Reinforcement learning for a visually-guided autonomous underwater vehicle, *in Proceedings of the International Symposium on Unmanned Untethered Submersibles Technology (UUST'99)*, Durham, New Hampshire, USA.

Contents

List of Figures

List of Tables